

Literature review methodology for scientific and information management, through its structuring and systematization

Metodología para la revisión bibliográfica y la gestión de información de temas científicos, a través de su estructuración y sistematización

Eduardo Gómez-Luna ^a, Diego Fernando-Navas ^b, Guillermo Aponte-Mayor ^c
& Luis Andrés Betancourt-Buitrago ^d

^a Dr. Ing Electricista. Grupo GRALTA, Universidad del Valle, Colombia eduardo.gomez@correounivalle.edu.co

^b M.Sc, Ing. Electricista. Profesor Asistente, Grupo GRALTA, Universidad del Valle, Colombia diego.navas@correounivalle.edu.co

^c Dr. Ing. Electricista. Profesor Titular, Grupo GRALTA, Universidad del Valle, Colombia guillermo.aponte@correounivalle.edu.co

^d M.Sc, Ing. Grupo GAOX, Universidad del Valle, Colombia luisandresbetancourt@gmail.com

Received: February 11th, 2013. Received in revised form: November 13th, 2014. Accepted: February, 18th, 2014

Abstract

This paper presents a methodology to find, organize and analyze the information in any field of research, thereby facilitating the acquisition of the available information and the identification of the main authors, the number of publications per year, the main areas of work and the future trends in a proposal topic. It is also, shows an application of the methodology for the topic "Frequency Response Analysis with the transformer in service (FRA on-line)".

Keywords: Literature review, database, technology watch, bibliometric.

Resumen

Este documento presenta una metodología para búsqueda, organización y análisis de la documentación en cualquier campo de investigación, facilitando la adquisición de la información disponible y la identificación de los principales autores, el número de publicaciones por año, las principales áreas de trabajo y las tendencias futuras de un tema propuesto. Se presenta también una aplicación de la metodología para el tema de "Análisis de la respuesta en frecuencia con el transformador energizado (FRA on-line)".

Palabras Clave: Revisión bibliográfica, base de datos, vigilancia tecnológica, gestores de referencias, bibliometría.

1. Introducción

El trabajo de revisión bibliográfica constituye una etapa fundamental de todo proyecto de investigación y debe garantizar la obtención de la información más relevante en el campo de estudio, de un universo de documentos que puede ser muy extenso [1]. Dado que en la actualidad se dispone de mucha información científica y su crecimiento es exponencial, el problema de investigar es precedido por el ¿cómo? manejar tanta información de forma eficiente [2].

La revisión bibliográfica o estado del arte corresponde a la descripción detallada de cierto tema o tecnología, pero no incluye la identificación de tendencias que puedan plantear diferentes escenarios sobre el desarrollo de la tecnología en cuestión y que permitan tomar decisiones estratégicas. Para ello se emplea la Vigilancia Tecnológica [3-4], la cual permite comprender los cambios tecnológicos

del entorno a través de la revisión periódica de diferentes fuentes. Este último aspecto no es tema del presente artículo.

La metodología que se presenta para la revisión bibliográfica está compuesta de tres fases fundamentales, las cuales deben retroalimentarse a lo largo de la búsqueda.

2. Paquetes computacionales de apoyo

En general, existen muchas plataformas en forma de software que permiten realizar estructuración de la información y en ciertos casos, algunas son tan avanzadas que hacen minería de datos y estadística multivariada para facilitar la obtención de información que no se tendría a simple vista.

Algunos de estos programas se muestran en la Tabla 1.

Tabla 1.
Software de apoyo para la gestión documental

Nombre	Características	Ventajas
GOLDFIRE Apoyo a la VT y gestión documental	Es un software analizador que lee los contenidos de Internet, de bases de datos locales, de bases de artículos científicos y de patentes.	Identifica las soluciones a preguntas hechas en lenguaje natural, permite la extracción de valor de los documentos. Permite establecer la evolución de una tecnología, el perfil de una organización en cuanto a su actividad de patentes; identificar las organizaciones implicadas en el desarrollo de determinadas patentes, la citación de patentes, de inventores, etc.
WebQL	Es una herramienta de software diseñada para recuperar información de fuentes no estructuradas, semi-estructuradas, y estructuradas.	La información es extraída en diferentes formatos: HTML, XML, pdf, doc, CSV, TSV, imágenes, bases de datos, etc. También, permite navegar a través de sitios pertenecientes a la “deep web”.
MATHEO PATENT Apoyo a VT solo de patentes y gestión documental	Es un software diseñado para automatizar las tareas de búsqueda, recuperación y análisis de patentes de las bases de datos de la Oficina de Patentes de los EE.UU. (USPTO) y de la Oficina Europea de Patentes (EPO).	Permite la búsqueda automática de patentes, y su gestión documental.
Biowiadom (antes Refviz u Omniviz)	Es un software de visualización de documentos científicos o estructurados.	Permite hacer agrupación de información basado en técnicas estadísticas multivariadas para determinar similitud de documentos que se traducen en áreas de trabajo o temáticas.
Zotero	Es un software para la gestión documental en la investigación. Se basa en la organización categórica de los documentos y referencias en múltiples formatos.	La ventaja de Zotero está en su integración con los navegadores de Internet, la posibilidad de sincronización y el uso de diferentes formatos.
Mendeley Desktop	Al igual que Zotero es un gestor documental de información científica. Está más orientado a ser una red social científica.	Las ventajas están en su utilización intuitiva, el manejo de documentos, la facilidad de importación y exportación y el compartimiento de documentos en grupos privados y públicos, dando soporte al trabajo de investigación en la red social científica.
JabRef	Es un gestor documental de código abierto desarrollado sobre Java.	Es una multiplataforma de código abierto, que opera bajo el formato BibText, lo cual permite la personalización y adaptabilidad a las distintas fuentes de información. Permite crear etiquetas automáticas por títulos, autores y otros campos adicionales.

3. Metodología

La metodología propuesta para la revisión bibliográfica puede ser aplicada a cualquier tema de investigación para determinar la relevancia e importancia del mismo y asegurar la originalidad de una investigación. Además, permite que otros investigadores consulten las fuentes bibliográficas citadas, pudiendo entender y quizá continuar el trabajo realizado. La metodología propuesta se compone de tres fases:

3.1. Definición del problema

Debe ser lo suficientemente clara para poder realizar una búsqueda bibliográfica que responda a las necesidades del investigador en particular, y que además aporte al estado de la técnica, de manera que conduzca a un escenario bastante amplio y permita la retroalimentación de la investigación.

3.2. Búsqueda de la información

Para el proceso de investigación bibliográfica se debe contar con material informativo como libros, revistas de divulgación o de investigación científica, sitios Web y demás información necesaria para iniciar la búsqueda.

Una búsqueda bibliográfica debe hacerse desde una perspectiva estructurada y profesional. Leer documentación

que no tenga fundamentos es aburrido y termina por ser una pérdida de tiempo. Por supuesto, cuando se inicia un proceso de búsqueda bibliográfica no se sabe qué material es el más pertinente o relevante, sin embargo, a medida que se avanza la perspectiva mejora y se empiezan a definir los temas que realmente interesan. Por tanto, se hace necesario delimitar la búsqueda y saber cuándo parar, aunque exista un sin número de preguntas por responder antes de abordar el tema principal del proyecto.

El material que se emplee debe ser “reconocido”, es decir, no puede consistir meramente en conversaciones de pasillo o en archivos descargados de Internet sin mayores referencias. Los trabajos reconocidos son aquellos que han sido revisados cuidadosamente por expertos antes de ser publicados.

La información se presenta en multitud de formatos, unos más accesibles, más reconocidos y mejor valorados “académicamente” que otros, entre estos: libros, revistas, actas de congresos, reportes técnicos, normas, tesis e Internet. Los libros son el punto de partida de cualquier investigación bibliográfica, pues proporcionan una buena base y una visión global del tema elegido [5].

Las revistas contienen artículos (generalmente revisados por expertos) actualizados y especializados, por lo tanto, más difíciles de asimilar. En el caso que no se entienda

alguna teoría o concepto que no deja avanzar en el proyecto, lo más pertinente es estudiar los libros para adquirir una base sólida antes de intentar entender los artículos más actuales y las últimas teorías publicadas en revistas.

Las actas de congresos contienen artículos presentados en conferencias nacionales e internacionales. La calidad de estos artículos varía mucho, dependiendo del proceso de revisión y de la experiencia y trayectoria de los revisores.

Los reportes técnicos por lo general son informes de proyectos de investigación realizados por universidades y/o empresas públicas o privadas, en los que se encuentra información de gran utilidad. En ocasiones es difícil acceder a estos ya que la política de publicación depende de la entidad que los realiza, en algunos casos los reportes técnicos dan información para alguna aplicación en particular y promover cierta marca o producto.

Las normas ó estándares a nivel nacional o internacional, marcan una pauta o guía a considerar en la ejecución de procedimientos específicos. Son el resultado de las opiniones y sugerencias de expertos en un tema, permitiendo al investigador trabajar con fundamento y credibilidad.

Las tesis son las memorias e informes publicados como resultados de trabajos de investigación de maestría y/o doctorados, estas constituyen una fuente de ideas de actualidad y contienen una serie de valiosas referencias bibliográficas.

El Internet aparece como una valiosa fuente de información, pero se debe tratar con extrema cautela, pues se pueden perder horas navegando sin encontrar nada de valor y el material encontrado puede ser poco fiable.

Con el fin de definir el dominio de la investigación, se deben utilizar ecuaciones de búsqueda (expresiones que consisten de palabras claves y operadores lógicos) que incluyan términos específicos del tema a investigar, además, se deben estudiar los patrones semánticos y de citación e identificar en primera instancia la estructura cognitiva para así, determinar las principales líneas de investigación desarrolladas en el mundo para el estudio del tema. Se deben aplicar criterios de selectividad de manera que el investigador se enfoque en los documentos relevantes para la investigación y si, a pesar de todo, existe demasiada información, quizás el tema elegido sea excesivamente amplio y sea necesario enfocararlo aún más.

3.3. Organización de la información

Esta fase es de gran importancia en todo proyecto de investigación, consiste en organizar de manera sistemática la documentación encontrada. Se puede realizar tanto de manera básica o detallada. Inicialmente la información puede ser ordenada en carpetas u hojas de cálculo desarrolladas por el propio investigador de forma manual, sin embargo, el proceso es lento y deficiente; otra manera de hacerlo es mediante el uso de programas especiales tales como [6]: JabRef, Zotero, Mendeley, Endnote y Reference manager, mostrados en la Tabla 1. Estas aplicaciones son eficientes, de uso libre y permiten organizar fácilmente la información por título, autor, revista y aporte; además, generan fácilmente la bibliografía para el informe final.

Una manera de organizar la información es por relevancia, distinguiendo los principales documentos de los secundarios. Así se obtiene una estructura o diagrama que permite identificar los pilares del tema bajo estudio. Es necesario definir una estructura para organizar la información de forma jerárquica y la cantidad de datos que se van a incluir en esta (autores, año, resumen, idea principal, etc.).

Los mapas de ideas permiten relacionar los elementos más importantes con los niveles de cercanía que hay entre estos, para aclarar la visión que se tiene de un tema. En su ejecución se pueden combinar colores y formas, a fin de facilitar la claridad visual del contenido.

Se pueden usar diagramas jerárquicos (esquemas circulares, red de contenido, esquema de barras, esquema de sangrado, esquema de llaves, etc.) para diferenciar claramente los documentos principales de los secundarios. Su elaboración requiere del conocimiento de las ideas centrales y los postulados fundamentales del tema.

Otra forma de organizar la información es mediante el uso de tablas de datos, cuyo fin es tener una estructura organizativa determinada de los datos y hacer una presentación concisa de los mismos. Las variables o características se limitan a los criterios escogidos por el autor.

3.4. Análisis de la información

La tercera fase es analizar la información ya organizada, indagando sobre cuáles son los documentos más útiles para la temática en estudio.

El análisis de la información es la tarea que toma más tiempo en la investigación bibliográfica, ya que con ella se espera identificar el aporte a realizar. En esta fase se debe tener un pensamiento crítico y debe ser realizada en paralelo con la primera, dado que es un proceso constante. Es un ciclo donde se reafirman las ideas planteadas en la formulación del problema y si se conoce bien el problema, la solución estará al alcance. Un ejemplo de las diferentes estrategias utilizadas para un adecuado análisis de la información se ha aplicado en diferentes estudios donde los resultados han sido satisfactorios. [7-8]

Una herramienta informática que facilita el análisis cualitativo de un grupo de documentos es *Refvix*, mediante la cual se puede obtener un mapa de proximidad de los artículos de acuerdo con la frecuencia de aparición de palabras en el texto, así como la creación de los listados de descriptores primarios, identificados en los grupos temáticos más significativos.

Una vez identificados los documentos principales, es necesario realizar un análisis de co-citación de autores y de co-ocurrencia, para identificar los frentes de investigación y los autores con más citaciones [9-10]. También se debe realizar un análisis detallado de los artículos que se identifiquen como directamente relacionados con las ideas más importantes y los aspectos relevantes para el tema de estudio.

El estudio de los documentos más importantes posibilita el reconocimiento de ideas principales, inferencias, conceptos claves, etc.

Se puede además realizar un análisis cuantitativo, en el que se estudien aspectos como número de documentos por año, documentos, citas por autor, documentos e investigaciones realizadas por país, etc.

4. Aplicación de la metodología a un tema de investigación en ingeniería eléctrica

La metodología propuesta se aplicó a un tema de ingeniería, relacionado con transformadores eléctricos, siguiendo las etapas indicadas antes:

4.1. Definición del problema

Como tema se definió el “análisis de la respuesta en frecuencia con el transformador en servicio (FRA on-line)”.

4.2. Búsqueda de la información

Una vez definido el problema a investigar, se consultaron diferentes fuentes de información, entre normas, reportes técnicos, patentes, revistas especializadas, memorias de conferencias y simposios y tesis doctorales; empleando las ecuaciones de búsqueda mostrada en la Tabla 2.

Tabla 2.
Ecuaciones de búsqueda empleadas para el caso de investigación.

Ecuación de búsqueda	Resultados
Transfer function AND transformer	522
Frequency Response Analysis AND transformers	592
FRA AND transformer	86
monitoring transformer winding AND frequency response analysis	31
Monitoring transformer AND transfer function	44
Diagnosis in transformer AND transfer function	44
Monitoring AND transformer AND FRA	30
Monitoring on line in transformers AND transfer function	17
Monitoring on line in transformers AND Frequency Response Analysis	16
Total	1382

La Figura 1 muestra los porcentajes de impacto de cada fuente sobre el total de la búsqueda realizada, lo cual define el dominio de la investigación.

Como se observa en la Tabla 2, se emplearon nueve (9) ecuaciones de búsqueda, teniéndose entre algunas de ellas solo leves modificaciones, pero que alteran los resultados encontrados. Es por ello importante en este paso considerar todas aquellas palabras claves que estén relacionados con la temática bajo estudio, incluso las abreviaciones empleadas, para reducir la probabilidad de pérdida de información valiosa.

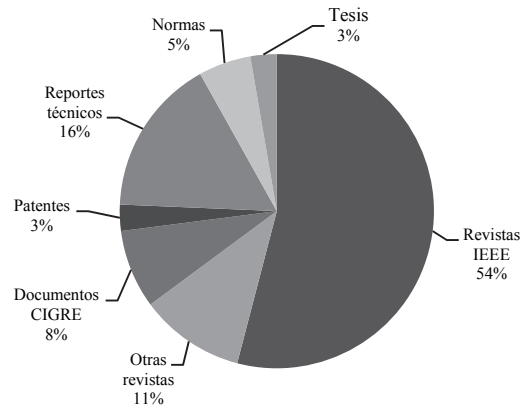


Figura 1. Material informativo consultado durante la búsqueda.

4.3. Organización de la información

Los 1382 documentos encontrados fueron organizados y analizados, aplicando los criterios de selectividad para identificar los más relevantes; para lo cual se empleó el software *Refviz*, el cual agrupa por temáticas y de forma autónoma (ecuaciones de búsqueda) toda la documentación. Esto permite ver todo el universo de búsqueda clasificado en temáticas, en un mapa de ideas que el software denomina una galaxia (Figura 2).

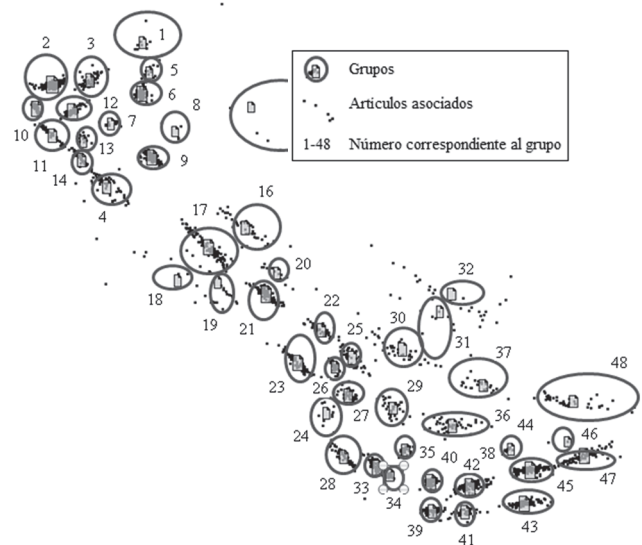


Figura 2. Organización sistemática de la documentación a través de un mapa de ideas (galaxia).

Los 1382 documentos encontrados inicialmente fueron agrupados en 48 temas de interés, mostrados en el mapa de ideas o galaxia de la Figura 2, donde cada idea se encierra en un círculo, que a su vez representa uno de los grupos encontrados (tema) con los documentos relacionados, facilitando el análisis de la información en las etapas posteriores.

Adicionalmente a *Refviz*, se empleó *Mendeley* para organizar por categorías los documentos. En la Figura 3 se observa un pantallazo con parte de los resultados para el caso objeto de estudio.

Esto permitió la creación de los diagramas jerárquicos

(esquemas de barras) mostrados en las Figuras 4, 5 y 6 que ayudan a diferenciar los documentos principales de los secundarios.

En estas Figuras se puede visualizar aspectos como: los autores que más han publicado en la temática de estudio, los que más han sido citados y la relación de las publicaciones

con los años.

La Figura 7 muestra una tabla de datos donde se presentan los diferentes países en los que se han realizado investigaciones en el tema, lo cual puede ayudar en la búsqueda de posibles aliados investigativos.

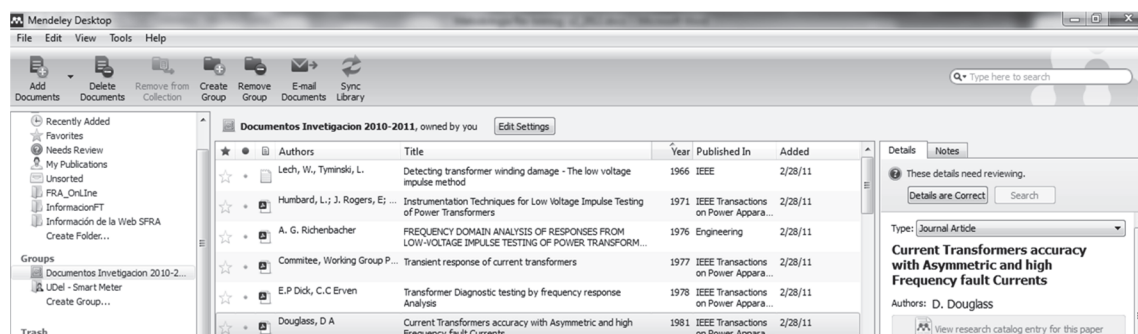


Figura 3. Organización de la información mediante la herramienta Mendeley.

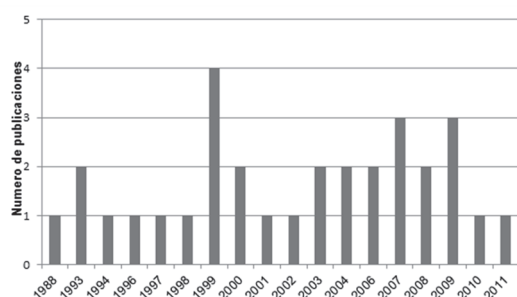


Figura 4. Número de publicaciones por año.

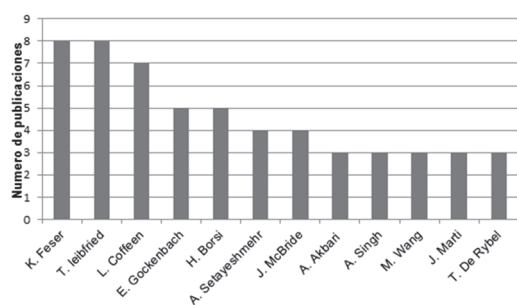


Figura 5. Número de publicaciones por autor.

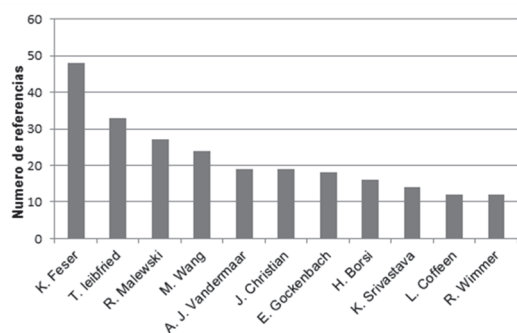


Figura 6. Autores más citados en los documentos .

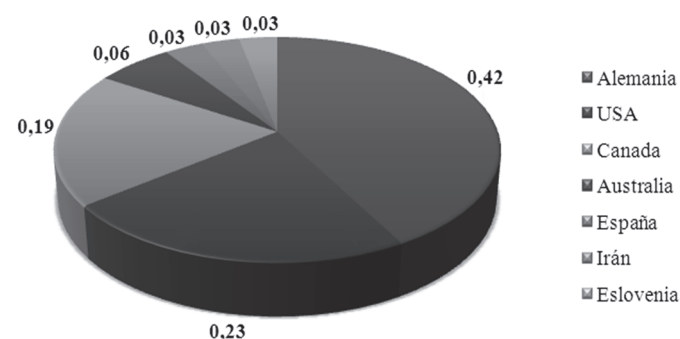


Figura 7. Países en los que se publicaron los documentos.

Mediante estos diagramas, se logra tener una idea más clara del tema y se empieza a focalizar sobre los aspectos de particular interés.

4.4. Análisis de la información

La información organizada, se debe analizar, para lo que se empleó la herramienta Mendeley. Identificados los autores con más citaciones (Figura 6) se realizó un análisis sobre los principales artículos con las ideas más importantes y los aspectos más relevantes para el tema de estudio, mediante la lectura de los resúmenes y conclusiones de dichos artículos.

Una vez realizado este trabajo, de los 48 grupos conformados en la etapa de búsqueda de la información, se seleccionaron 8 como los de mayor interés, los cuales son mostrados en la Figura 8. Estos grupos los conforman 71 artículos.

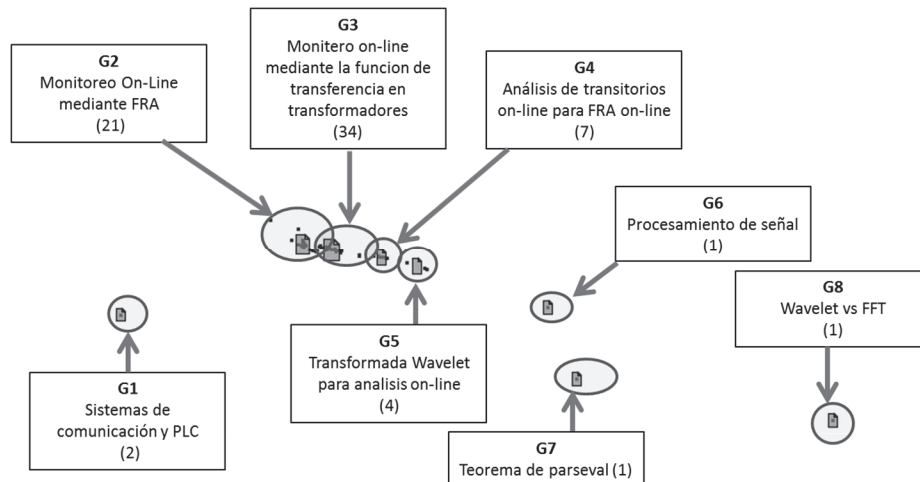


Figura 8. Grupos de interés para la investigación.

En esta etapa de la metodología se observa un avance significativo, ya que de los 1382 documentos iniciales se tienen 71 identificados como de mayor interés (primarios).

Los documentos filtrados debieron ser leídos con mayor detalle, para de esa manera finalmente seleccionar los de verdadera relevancia para la investigación. La metodología concentra la lectura sobre un menor número de artículos ya identificados como de mayor interés. Sería imposible e impráctico leer detalladamente dentro de un tiempo prudente los más de 1000 documentos iniciales.

Con la lectura detallada, finalmente se seleccionaron 23 artículos de los 71 y con estos se elaboró el estado del arte sobre el tema la respuesta en frecuencia con el transformador en servicio (FRA on-line), el cual fue publicado en una revista internacional [11].

5. Conclusiones

En este artículo se presentó una metodología para realizar revisión bibliográfica sobre cualquier tema de investigación, mediante una macro búsqueda que permite la identificación de los documentos referentes al tema de investigación.

Las estrategias de búsqueda, organización y análisis de la información, permiten tanto la obtención de los documentos referentes a un tema de investigación, así como su sistematización y estructuración con el objeto de analizar las principales características del conjunto de documentos bajo estudio.

Se presentó un caso de aplicación sobre el tema de análisis de la respuesta en frecuencia de transformadores energizados (FRA on-line), en donde se identificaron las principales categorías y áreas de interés, y se realizó la organización y el análisis de los documentos a través de Mendeley y Refvz respectivamente.

La metodología planteada permite realizar de manera organizada y secuencial una de las etapas más importante de todo proyecto de investigación: “la revisión bibliográfica”.

Referencias

- [1] Aleixandre-Benavent, R, González Alcaide, G, González De Dios, J, y Alonso-Arroyo, A. Fuente de información bibliográfica (I). Fundamentos para la realización de búsquedas bibliográficas. *Acta Pediátrica Española*, 69 (3), pp. 131-136, 2011.
- [2] Vilanova, J.C. Revisión bibliográfica del tema de estudio de un proyecto de investigación. *Radiología*., 54 (2), pp. 108-114, 2012.
- [3] Casteñanos, O. F. Torres, L.M. y Rossero, J. Aplicación de un modelo de inteligencia para la definición de estrategia tecnológica en diferentes niveles de complejidad institucional, en *Seminario Latino-Iberoamericano de gestión tecnológica* (XI, 2005, Brasil). Memorias.
- [4] Rodríguez Salvador - Marisela. La inteligencia tecnológica: elaboración de mapas tecnológicos para la identificación de líneas recientes de investigación en materiales avanzados y sinterización, tesis (Doctorado en Administración y Dirección de Empresas), España, Universidad Politécnica De Cataluña, 1999, 180 p.
- [5] Cisneros, M y Olave, G. Redacción y publicación de artículos científicos: enfoque discursivo, 1a ed., Bogotá, Editorial ECOE, 2012, pp. 139.
- [6] Rivera, F. y García Rojo, M. La búsqueda de bibliografía: nuevas soluciones para un viejo problema. *Nefrología*., 23 (6), pp. 487-496, 2003.
- [7] Rojas-Sola, J. I., Fernández-Sora, A., Serrano-Tierz, A Y Hernández-Díaz, D. Una revisión histórica: desde el dibujo en ingeniería hacia la ingeniería del diseño, *Revista Dyna*., 78 (167), pp. 17-26, 2011.
- [8] Bedoya, A., Castrillón, J. C., Ramírez, J. E., Vázquez, J. E Y Zabala, M. A. Producción biológica de hidrógeno: una aproximación al estado del arte, *Revista Dyna*., 75 (154), pp. 137-157, 2008.
- [9] White, H. D And McCain, K. W. Visualizing a discipline: an author co-citation analysis of information science, *Journal of the American Society of Information Sciences*., 49 (4), pp. 327-355, 1998.
- [10] Salton G. Automatic indexing using bibliographic citations, *Journal of Documentation*., 27 (2), pp. 98-110, 1971.
- [11] Gómez-Luna, E., Aponte, G., González-García, C. And Pleite, J. Current Status and Future Trends in the Frequency Response Analysis (FRA) with the Transformer in Service, *IEEE Transactions on Power Delivery*., 28 (2), pp. 1024-1031, 2013.